Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

may be more potent and significant than others; and, finally, some may be recurring and regular while others are occasional and accidental. That several kinds of fluctuations appear in any time series is obvious. For example, we expect pig-iron production to be vastly greater in 1910 than in 1890, because of the growth of the industry in the interval. On the other hand, we know from experience that there are periods of prosperity and depression, and are not surprised, notwithstanding the large increase *for the whole period*, to find considerable ups and downs in the production of pig iron *during the period*. Irregular or accidental occurrences, likewise, cause fluctuations — a strike, a court decision, the outbreak of war. Within any given year, also, there are fluctuations. The round of the seasons is known to produce changes in the volume of production or of sales, as well as other related changes. Moreover, *all* of these forces may be at work simultaneously: and all of them play a part in determining the fluctuations which occur in a statistical series. Those fluctuations are thus a confused conglomerate growing out of numerous causes which overlie and obscure one another.

In view of this complexity the simple comparison of items within time series, in what may be termed their crude state, is of but little significance. How ascertain from such a conglomerate the particular meaning of any given change in the figures ? How divide the given change into its constituent parts ? In other words, how isolete and measure the influence of different sets of ses or forces ? These are the difficult problems which render necessary the development of an adequate method for the analysis and comparison of statistics of business conditions.

Stated positively, the problem is to devise a method of unravelling, in so far as may be, the tangle of elements which constitute the fluctuations in fundamental series; that is, of distinguishing and measuring types of changes; of eliminating certain types; and thus of revealing in the corrected data the quantitative effect of those forces which underlie the business and industrial situation. In other words, the problem is to devise a method of analysis whereby the crude time series of business statistics now published in our trade and financial journals may be set forth in such a form as to be significant and reliable indices of business conditions.

It is essential, therefore, that at the outset of this study there should be a clear-cut definition of the assumptions which we have made concerning the nature of the fluctuations found in time series. A preliminary survey of the graphs of such fundamental monthly series as bank clearings, iron production, commodity prices, and new building permits has led to the following working hypothesis, namely that each series is a composite consisting of four types of fluctuations. The four types are:

1. A long-time tendency or secular trend; in many series, such as bank clearings or production of commodities, this may be termed the growth element;
2. A wave-like or cyclical movement superimposed upon the secular trend; these waves appear to reach their crests during periods of industrial prosperity and their troughs during periods of industrial depression, their rise and fall constituting the business cycle;
3. A seasonal movement within the year with a characteristic shape for each series;
4. Residual variations due to developments which affect individual series, or to momentous occurrences, such as wars or national catastrophes, which affect a number of series simultaneously.

Although these four types of fluctuations occur in all series, they are by no means uniform in different series. Each series appears to have a characteristic secular trend for a selected period;[1] each appears to have cycles, but these cycles are neither synchronous nor of the same intensity;[2] each appears to have a seasonal movement and irregular variations peculiar to itself. In consequence of these individual traits each series must be treated by itself. General methods may, indeed, be developed; but such methods must always take into account the special characteristics of each series.

With this brief preliminary survey, we proceed to the separate consideration of the types of fluctuations.

2. MEASUREMENT OF SECULAR TREND

Current statistical items such as bank clearings, interest rates, prices, and number of business failures convey a meaning only in case we are able to compare such items with others of a similar nature for other periods, or with a standard of some sort. We require a " normal " as a bench mark from which to reckon. For statistical items pertaining to different time intervals this " normal " may not be constant. There may be a normal change year after year in a developing or changing industrial society, just as there is a normal change in the physical or mental status of a growing child. The initial problem to be solved in using statistics as indices of business conditions is the problem, first, of ascertaining if there is a change in the items year after year, second, of determining the normality or regularity of such change, and third, if there be a normal change, of measuring it and correcting the original items for it. This is the problem of the secular trend. Graphically, the problem of secular trend may be thought of as the problem of securing a line or curve, the general direction and position of which are determined by the data.[3]

Both the existence and the measurement of secular trend are to be established, not on *a priori* grounds, but by examination of the statistics themselves. To be sure, it is sometimes possible to deduce the *direction* of the movement of a given series in advance of special investigation. For instance, the trend of such series as the production of commodities in a growing country may be predicted as upward. But it is only when relationships are simple, and the change is very considerable, that such conclusions inspire any confidence. In any case, deductive reasoning does not supply quantitative results; and it is precisely such results that are requisite if we are to separate a composite into its elements. Clearly, we must work from the data themselves. Moreover, as has been pointed out in a previous place, we must work with each series separately. Though each series is a composite of the four types of fluctuations that have been named, in no two series are these fluctuations identical. As to the type of fluctuation now being considered, the secular trend, it is apparent from general observation that even the direction, to say nothing of the extent, of change may be different in different series. Wholesale prices of commodities, for instance, may move upward while interest rates move downward.

---

[1] See " The Variate Difference Method and Curve-Fitting," by W. M. Persons, *Quarterly Publications of the American Statistical Association*, June, 1917.

[2] See " The Construction of a Business Barometer; Annual Data," by W. M. Persons, *American Economic Review*, December, 1916.

[3] See Chart 1, *et seq.*

Finally, the secular trend must be measured for a given period, a period defined with reference to the specific problem which is being investigated. But there are puzzling questions attached to the definition of the period. In a study of statistics of economic conditions, the chief desideratum is to find a means of interpreting current statistical data. It is obvious that the meaning of current data must be based upon generalizations concerning the significance of such data in the past. At the basis of our judgment, therefore, is an assumption that the economic conditions of the period contributing the data for our generalizations are essentially homogeneous with the conditions of the present and immediate future. That such homogeneity exists, however, is by no means necessarily the case. To illustrate: the trend of general commodity prices for the period 1879–1896 was downward; from 1896 to the present it has been upward. For all other statistical series examined, whether the items be stated in terms of dollars or of physical units, the annual increase has been greater during the latter period than during the former. Looking at the past we conclude, therefore, that for these series the secular trend depends on the period taken. For the entire period 1879–96 it may be represented by a parabola — concave upward — with its vertex in the 90's, or by some other more complicated function.

That is, looking to the past, we have provided for a *change* in the direction of the secular trend, the change being due to the fact that the past period considered contains two unlike subperiods. But is it reasonable to assume that the function which pictures the general trend of the series in the past will also hold for the indefinite future ? In case of such series as iron production and bank clearings the assumption that the items in the future will tend to increase with a constant increment, either absolute or relative, may be reasonable on *a priori* grounds. For the trend of such series is apparently due, primarily, to the increase of population and the consequent growth of the industry. Or, on *a priori* grounds, we may expect a steady decrease in certain series, such as interest rates. But such assumptions appear not to be valid in cases where an industry (the automobile industry, for example) is changing radically in the position it occupies in our industrial life. Even less valid is an assumption of steady increase or decrease continuing into the future for series of commodity prices, security prices, interest rates, wages, and other series, the items of which are not functions of population. The secular trend satisfactorily determined for a past period does not always warrant us in extending the same trend into the present and future.

It is apparent from such considerations as these that the problem of computing secular trend raises some puzzling questions, one of the chief of which is the definition of the period for which the trend is to be computed. Generalizations from a past period are not applicable to the present or future unless the conditions of the past period are homogeneous with those of the present and future. That such homogeneity exists is not always easy to ascertain. Still more difficult is it to locate the precise point in a series at which significant changes in homogeneity occur. Sometimes, to be sure, the break in homogeneity may come about with suddenness, as the result of some catastrophe. More often there is no sharp break, but an elusive change which comes about imperceptibly, that is, by small increments. Small yearly increments affect the homogeneity of adjacent years very little, but their cumulative effect is so great as to render the later portion of a series of years distinctly unlike the earlier portion. Moreover, from whatever cause or in whatever manner the change may come about, it is difficult to ascertain *at the time* whether continuity with the past has been definitely interrupted, or whether the change is merely a short-time fluctuation of minor import so far as secular trend, and homogeneity essential thereto, is concerned.

Despite these difficulties in judging the past, however, we are obliged to appeal to the past for generalizations applicable to the present. The questions to be answered, then, are the following: — For how long a period is homogeneity to be expected ? How is homogeneity to be ascertained ? What degree of homogeneity is required ? It is to be expected that the *immediate* past is more homogeneous with the present than a more remote period. On the other hand, although recent tendencies should undoubtedly be given more weight than those more remote, a computation of secular trend on the basis of too brief a period would have no significance. Even considerable changes occurring in a series within two or three years provide no warrant for the conclusion that the secular trend has materially changed. For the secular trend is a " long-time " tendency. The question is, how long a time is necessary to show significant trend; how long a time can safely be taken consistent with the requirement of homogeneity of data ? To this question no absolute answer can be given. Consideration of the results of handling the data in various ways, however, establish the following criteria. We have first, as a broad guide, the fact that secular trend is by definition a " long-time " trend;

long, that is, with respect to the problem in hand. In a study of business conditions, therefore, it is a trend covering a period at least as long as a full business cycle of prosperity and depression. This indefinite period is the minimum for which it may be determined. The maximum period is fixed by the homogeneity of the data; but since there are no criteria for deciding precisely what degree of homogeneity is requisite for all series, this maximum also is indefinite. If it is so long as to include heterogeneous subperiods of widely different fundamental conditions, the problem of determining the secular trend is greatly complicated and its solution of questionable value. The maximum period to be taken must, therefore, be determined by examination of the statistical series used, and of the general conditions affecting the series.

With this brief discussion of the concept of secular trend, and of the implications and conditions which the concept involves, we may proceed more intelligently to the consideration of methods of measuring the trend. Two methods of measuring and eliminating secular trend were tested by applying them to various series. Both are empirical. The first method is based on moving averages covering a number of years; the second method is that of fitting lines or curves to the series for the period selected, by the application of mathematical formulas.

The steps in the application of the first method to annual data are: (*a*) The selection of a period for which the moving averages are to be computed. If we have a series of annual items beginning with 1879 and select ten years as the period for moving averages, we compute averages of the items for the ten years 1879–88, 1880–89, 1881–90, and so on. (*b*) The expression of the original annual items as percentages of the average for the preceding ten years. Thus, the item for 1889 would have as its base the average for the ten years 1879–88, the item for 1890 would have the average for 1880–89 as its base, and so forth. The resulting percentages are supposed to be corrected for secular trend.[1]

The question that immediately arises in connection with this method is this: what effect has the choice of the period of the moving average? To answer this question the method described was applied to the following series of annual items:

Per capita gross receipts of railroads of the United States (1879–1913).[2]
Per capita coal produced in the United States (1879–1913).[2]
Per capita exports of the United States (1879–1913).[2]
Per capita steel production of the United States (1879–1913).[2]
Per capita imports of the United States (1879–1913).[2]
Per capita bank clearings of New York City (1879–1916).
Gross bank clearings of New York City (1879–1916).
Gross bank clearings of the United States outside New York City (1879–1916).

Five sets of moving averages were then computed for each of the eight series, as follows: (1) five-year; (2) ten-year; (3) fifteen-year; (4) fifteen-year, weighted (the first five years were weighted 2; the second five years 3; the third five years 5); (5) twenty-year. From each of the original series five series of annual percentages or relatives were computed, the base for each series of relatives being the average for the preceding five, ten, fifteen, fifteen weighted, and twenty years, respectively. In other words, each relative is the percentage which the actual per capita figure for the year bears to the base. The resulting percentages were then plotted and compared. Chart A presents the results of the computations for two of the original series, per capita steel production and per capita gross receipts of railroads. The graphs for the other series are not reproduced because the conclusions to be drawn from them differ in no way from the conclusions drawn from Chart A.

Examination of Chart A revealed certain tendencies which were common to all the series and led to the following conclusions:

(*a*) The direction of the change from one year to the next (first difference) is the same in all cases, except that for gross receipts of railroads the five-year base gives percentages showing a contrary result to the others for the years 1896, 1897, and 1903. The magnitude of the changes from one year to the next depends to some extent on the base used. But since the object of expressing the items in terms of moving averages is to eliminate the secular trend from the relatives, observations concerning the correspondence of year-to-year fluctuations are not pertinent. The similarity of year-to-year fluctuations is mentioned here to show, first, that it has been noted, and second, that it has nothing to do with secular trend.

(*b*) The magnitude of the percentages themselves depends on the base used; the longer the base period the higher the percentages. It is obvious that this will be true for all series in which the original items show a steady upward tendency. The range of relatives in 1903–14 for the two series of Chart A is as follows:

	5-year base	10-year base	15-year weighted base	15-year base	20-year base
Gross Receipts of Railroads. .	104–122	113–140	118–142	127–150	136–156
Steel Product'n	70–143	84–179	91–190	104–218	122–254

Each series of percentages, then, has a plane of distribution which depends on the base taken. Before the

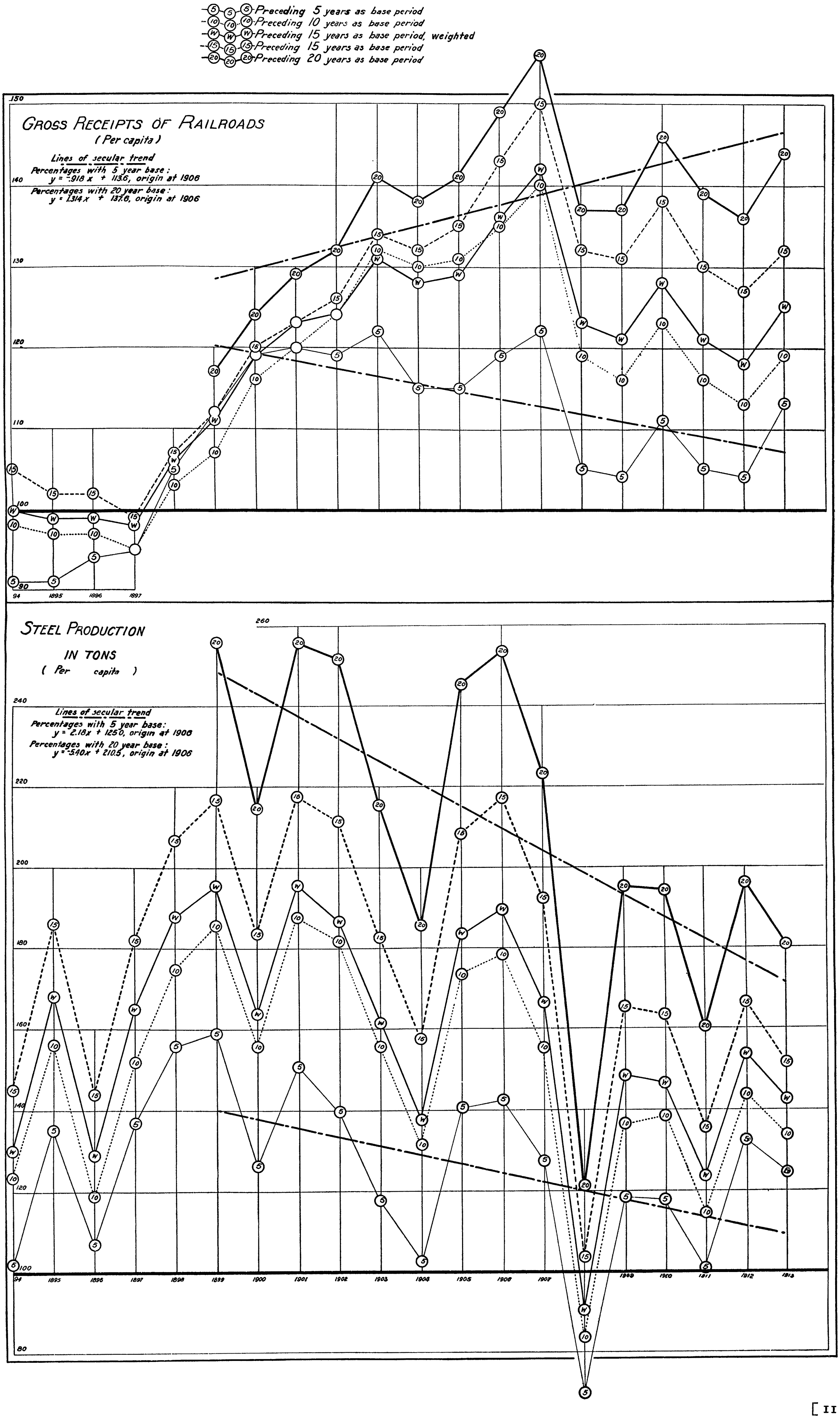
---

[1] It should be noted that the method of moving averages stops at the point indicated, that is, the expression of the items in terms of a moving base. No demonstration, or even analysis, of the efficacy of the method in actually removing trend from the data has heretofore been made. I have attempted to analyze, correct, and extend the method in the pages following.

[2] The per capita series were obtained by dividing the items of the corresponding gross series by the population of the United States as given in the *Statistical Abstract of the United States*.

CHART A

PERCENTAGE THAT ANNUAL ITEMS BEAR TO MOVING AVERAGE AS BASE

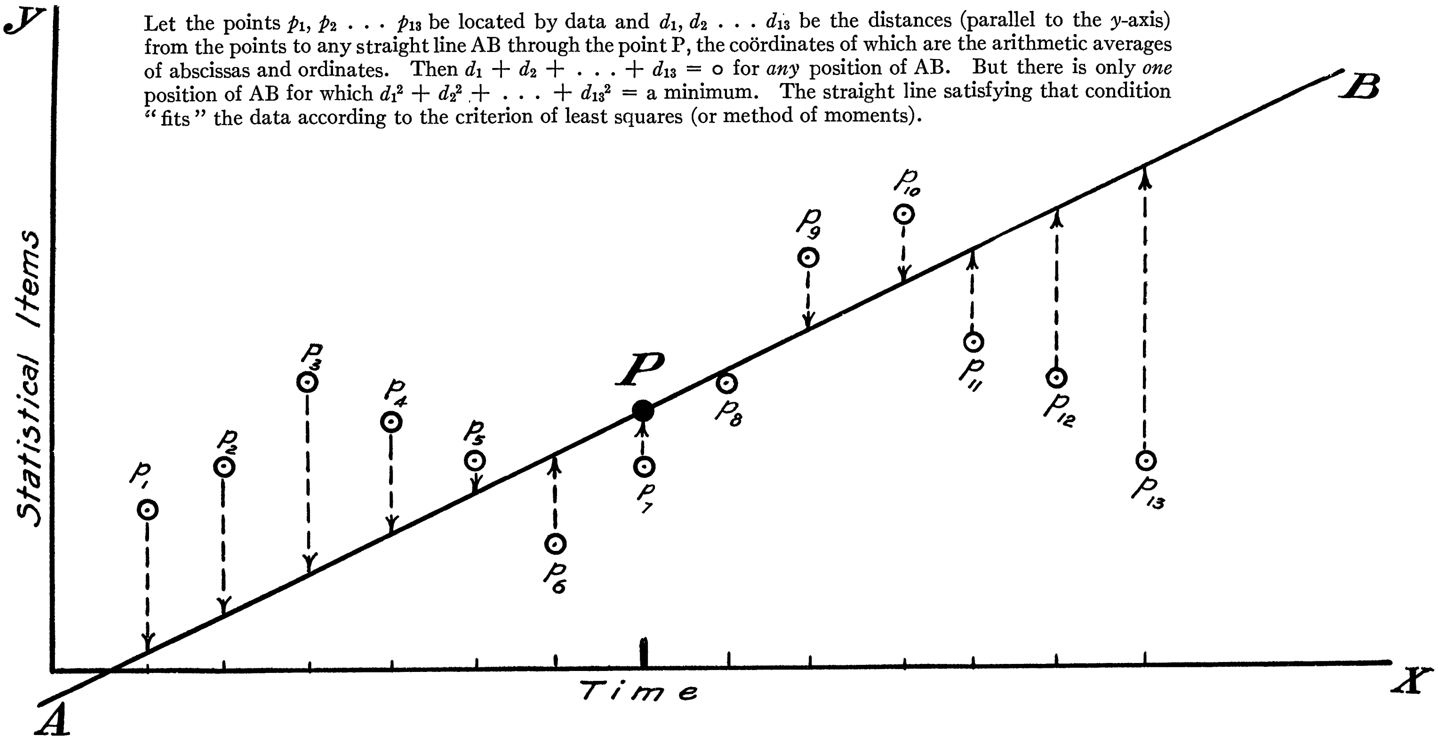


meaning of any percentage can be sensed this plane of distribution must be determined.

(*c*) The five series of relatives themselves exhibit five varieties of trend. The straight lines appearing in Chart A are fitted, by a method to be described later, to the relatives with five- and twenty-year bases, respectively, for the period 1900–14. These lines show that the relatives have a trend depending on the base used. The trend of the relatives for steel production is downward in all cases, the slope varying with the base selected. The trend of the relatives for gross receipts of railroads is downward where the five-year base is used, and upward for the twenty-year base.

It is clear from the preceding considerations that the use of moving averages as bases does not eliminate the secular trend of the original series. The resulting relatives present the problem with which we started, the measurement and elimination of the trend for the period in question.

The second method of eliminating secular trend is that of fitting a straight line or curve to the data. The process is as follows: (*a*) Plot the series for the period in the usual fashion with time as abscissas and the items as ordinates. (*b*) By inspection choose the type of function[1] that will best fit the data. (*c*) Using the data for the period in question, determine the constants of the chosen function by the method of least squares or method of moments. For a straight line (or a first degree function) these two methods of curve-fitting are identical. The straight line determined is the one such that the sum of the squares of the distances to it from the points corresponding to the items is a minimum.[2]



Let the points $p_1, p_2 \ldots p_{13}$ be located by data and $d_1, d_2 \ldots d_{13}$ be the distances (parallel to the *y*-axis) from the points to any straight line AB through the point P, the coördinates of which are the arithmetic averages of abscissas and ordinates. Then $d_1 + d_2 + \ldots + d_{13} = 0$ for *any* position of AB. But there is only *one* position of AB for which $d_1^2 + d_2^2 + \ldots + d_{13}^2 =$ a minimum. The straight line satisfying that condition "fits" the data according to the criterion of least squares (or method of moments).

---

[1] In mathematical terminology the word "function" refers to an equation expressing the relationship between two or more variables. Thus, the equations $y = mx + b$ and $y = ax^2 + bx + c$ are algebraic functions of y in terms of x. In those equations the letters a, b, c, and m represent constants for any known function. Such constants are termed "variable parameters" to indicate that different values of the constants give a "family" of curves. In other words, to each function with known parameters there corresponds a curve or locus; to each function with all possible parameters there corresponds a set of related curves.

[2] For the technique of least squares see Weld's *Theory of Errors and Least Squares*. The method of moments is treated by W. Palin Elderton in *Frequency-curves and Correlation*. The best treatment of correlation is given by Yule, *Introduction to the Theory of Statistics*, and Moore, *Forecasting the Yield and Price of Cotton*.

The straight lines determined by applying (*a*) the method of least squares, (*b*) the Pearsonian method of correlation, and (*c*) the method of moments are identical.

A demonstration of the identity of methods (*a*) and (*b*) may be conveniently found in Yule's *Theory of Statistics* (pp. 170–172). In the reference cited it is shown that the line of regression of Y on X found by correlating any time series of items (Y) with time (X) is also the line having a minimum sum of squares of distances from points located by the given items.

The line of regression (and the line of least squares) passes through the point located by the arithmetic averages, $\frac{\Sigma X}{n}$ and $\frac{\Sigma Y}{n}$. Since the coördinates $\left(\frac{\Sigma X}{n}, \frac{\Sigma Y}{n}\right)$ satisfy the first equation of condition, $\Sigma Y = m\Sigma X + nb$, of the method of moments it is clear that the line determined by that method also passes through the point of arithmetic averages. In consequence of this fact, the equations of the lines determined by all three methods may be put in the form $y = mx$ where x and y are deviations from their respective arithmetic averages or, in other words, where the origin is chosen at the point of arithmetic averages of the items and the corresponding time.

The slope, m, of the line of regression (also of the line of least squares) is $r\frac{\sigma_y}{\sigma_x}$ where r is the Pearsonian coefficient of correlation, σ_y is the standard deviation of the y series, and σ_x is the standard deviation of the natural numbers representing time. This expression for m reduces to the form $\frac{\Sigma xy}{\Sigma x^2}$.

The second equation of condition to be satisfied in the application of the method of moments is $\Sigma xy = m\Sigma x^2$. This condition gives the same value of m as that of the line of regression and of least squares. Therefore, the three methods give identical results.

Moreover, it may readily be shown that the line having the minimum sum of squares of the perpendicular distances from the given points also has the minimum sum of squares of distances parallel to the y-axis. If $d_1, d_2 \ldots d_n$ represent the last named distances we have $d = y - mx$ and, squaring and summing,

$$\Sigma d^2 = \Sigma y^2 - 2m\,\Sigma xy + m^2\,\Sigma x^2.$$

Differentiating the last named function with respect to m, equating the result to zero (condition for minimum), and solving for m we have, $m = \frac{\Sigma xy}{\Sigma x^2}$. This demonstrates the theorem stated at the beginning of the paragraph, since m is the value previously found.

It may also be shown that the algebraic sum of all deviations (parallel to the y-axis) from the given points to *any* line passing through this point of averages is zero. The demonstration of this

Let us consider the application of this method to the determination of the secular trend for the monthly pig-iron production of the United States, 1903–16. Chart 2 presents the actual data. Inspection of the data led to the conclusion that a straight line would "fit" the data for the period in question. It may be added that rarely, if ever, is it desirable to use a more complicated function in handling economic data. Therefore, a straight line was fitted to the annual totals by the method of moments.[1] The dot-dash line is this line of secular trend. On the same chart is shown the secular trend computed by the method of moving averages. The dotted curve below the straight line was located by the ten-year moving annual averages of the preceding period adjusted to months.[2]

Let us compare the trends obtained by the two methods. In order to make a comparison the method of moving averages will have to be extended and the percentages adjusted. As previously used by statisticians that method did not give an explicit measure of secular trend.

The production of pig iron increased, but not uniformly, during the period 1892–1916. The production for any given year is, therefore, nearly always *above* the average for the preceding ten years. The relatives having the average of the ten years preceding as their base are, on the whole, more than 100 per cent; on the basis of experience we would *normally* expect pig-iron production in the future to be above the preceding ten-year average. Even a depressed industry often produces in the current year a greater tonnage of pig iron than was produced as the average in the preceding ten years. The question is, in this particular case just what percentage above the ten-year average should we regard as normal. In considering this question, the 168 relatives were averaged, giving 128 per cent. We might assume, then, that a relative pig-iron production above 128 per cent represents an item above normal and a relative below 128 per cent, an item below normal. The dotted line shifted upward by 28 per cent of itself, say 28 per cent of 1600, or 450, is a graphic representation of the "normal" obtained by the use of moving averages. We have, then, two lines on our chart, each of which is an attempt to represent the normal. An examination of the chart shows that the two correspond reasonably well in this case. The straight line fitted to the data is somewhat higher at the beginning of the period, and is lower at the close than the line of moving averages, the latter moved upward 28 per cent.

Although the two methods give approximately the same results in this case, however, there is no question as to which is better. The method of fitting a straight line or curve is much superior to the method of moving averages. The latter furnishes an explicit measure only by application of a cumbersome adjustment that might not always prove satisfactory. The former is more direct and easier of application. It is much more reliable since it utilizes data only for the period under examination. That is, in curve-fitting, the data for the period 1903–16 constitute the raw material which determines the line or curve of secular trend for that period; whereas in case a ten-year moving base is used, the data for the ten years 1893–1902 are also required. The accuracy with which the straight line represents the trend becomes quite apparent upon inspection of the graph.

In measuring the secular trend for the period 1903–16, straight lines have been fitted by the method of moments to all the series examined. Examination of Charts 1 to 15 shows that the trend for this period is well represented by the straight lines found. In this connection, however, there remains to be considered

---

theorem comes readily from the consideration that the items of a time series are equal intervals apart so that the arithmetic average of the time is also the median. As the line through the point of averages (line *APB* in the accompanying figure) rotates about that point (*P*) the consecutive deviations to the right of the point are increased by the same amounts that the consecutive deviations to the left are decreased. Consequently, the algebraic sum of the deviations is a constant. Since the sum of the deviations is zero when the line *APB* is parallel to the x-axis the constant is zero.

There is but one line, however, passing through the point of averages such that the sum of the squares of the distances to the given points is a minimum. In other words, the line of least squares is uniquely determined. The accompanying figure illustrates the argument just presented.

[1] If y represents the annual items, n the number of items, and x the time in years measured from any convenient origin, and $y = mx + b$, the equation of the line to be fitted to the data, then the following equations are solved simultaneously for m and b:

$$\Sigma y = m\Sigma x + nb$$
$$\Sigma xy = m\Sigma x^2 + b\Sigma x$$

where the summation sign, Σ, indicates the sum of the items for the period taken. The constants m and b are, of course, easily reducible to a monthly basis. In the case of the series now under consideration, pig-iron production:

The sum of the annual totals, 1903–16, $= \Sigma y = 349{,}354$.

The number of annual totals, 1903–16, $= n = 14$.

We may measure time from any point as origin. To fix the origin in the middle of the period will make the algebraic sum of the abscissas, or $\Sigma x, = 0$. Where n is even, as in the present case, one unit of abscissa is made to represent a time interval of six months. In this case

$$\Sigma x^2 = (-13)^2 + (-11)^2 + \ldots + (-3)^2 + (-1)^2 + (+1)^2 + (+3)^2 + \ldots + (+13)^2 = 910.$$

Substituting these values in the equations, we have:

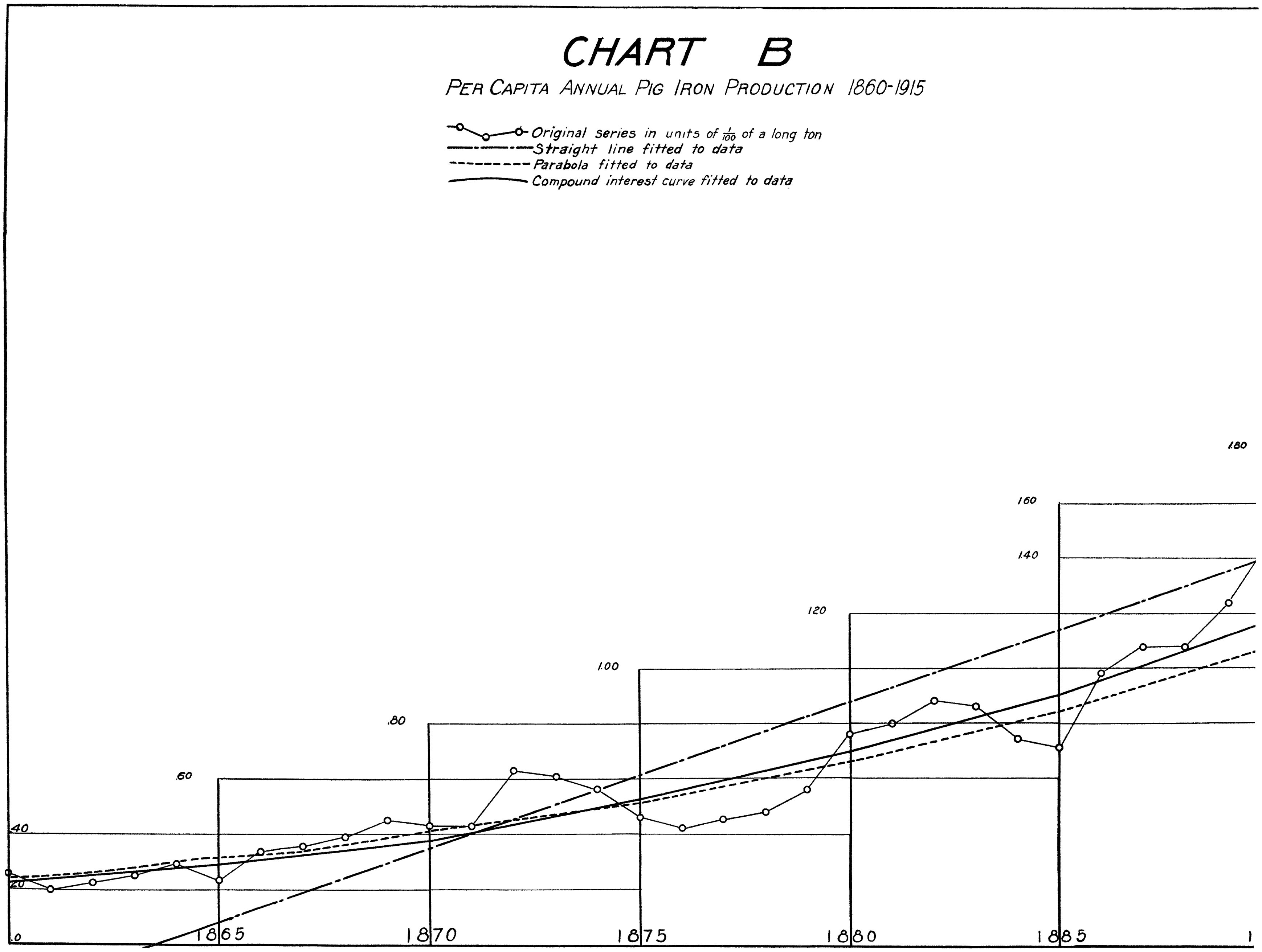
$b = \frac{\Sigma y}{n} = \frac{349{,}354}{14}$ for annual data or $\frac{b}{12} = \frac{349{,}354}{12 \times 14} = 2079$ for monthly data. (b is the ordinate of the secular trend midway between 1909 and 1910 measured in units of 1000 gross tons.)

$$m = \frac{\Sigma xy}{\Sigma x^2} = +\frac{519{,}484}{910} = 571 \text{ for annual data.}$$

(m is the uniform increment each six months.)

The uniform monthly increment is $\frac{m}{72} = \frac{571}{72} = 7.93$. Using the values $\frac{b}{12}$ and $\frac{m}{72}$ the various monthly ordinates of the secular trend are readily computed. The equation of the line of secular trend with the abscissa, x, measured in years from the middle of 1903, and the ordinate, y, measured in units of 1000 gross tons of pig iron produced per month, is $y = 95x + 1461$.

[2] Thus the ordinate of the dotted line for January 1903 is the ten-year total, 1893–1902, divided by 120; the ordinate for January 1904 is the ten-year total, 1894–1903, divided by 120; the items for the months of the year other than January are interpolated. Since monthly data of pig-iron production were not available previous to 1903 the ten-year averages are based on the annual data given in the *Statistical Abstract*. These data do not exactly coincide with those given by the *Iron Age* for the years 1903–16. However, comparison shows the discrepancy is negligible for the present purpose.



an important question. The straight lines do, indeed, fit the data for the *past* period, 1903–16, but do they constitute a good basis for estimating *future* trend? We may go very widely astray in extending the straight lines, and thereby estimating the trend for the new data, as they come in, to be the same as the trend for the period 1903–16. That is to say, the *arbitrary* extension into the present and future of the lines of secular trend determined for a past period, is a proceeding apt to lead to serious error. The current data should, in general, influence the position of the line or curve of secular trend. In other words, the straight line of secular trend should be adjusted by the utilization of current data as they come to hand.

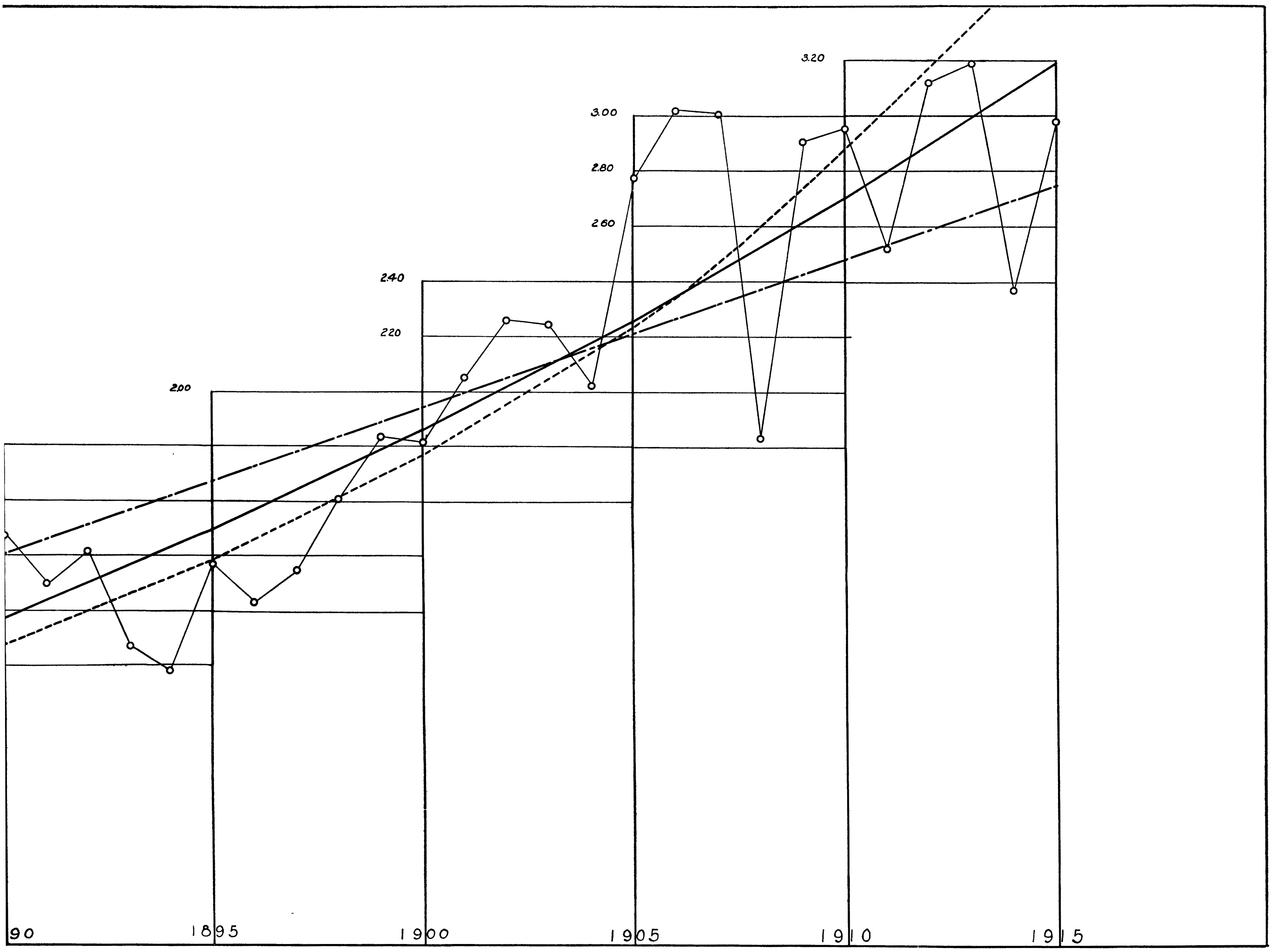
Our problem, then, is to ascertain by experiment the effect of using the straight line method of determining the secular trend, not for a completed past period, but for a changing period increasing in length through the addition of new data. The series chosen for test were annual wholesale price indices, 1879–1913, monthly total bank clearings, 1890–1916, and monthly average price of forty common transportation stocks, 1890–1915.[1] In all these series a decided change in trend occurred in the middle 90's; and they therefore present very perplexing problems. The series of annual commodity prices will be considered first.

In experimenting with the series named it was assumed that straight lines were fitted to the data of any period, say 1879–88, as though the process of fitting had taken place at the end of the terminal year, 1888, that is, with no information concerning developments after that year. In other words, the data were handled by a uniform method throughout; the method was not altered to correspond with the (now) known change of trend occurring in the 90's.

There is a hypothesis, resulting in a slightly altered operation of fitting a straight line to the data, that seems worth while testing. That hypothesis is that, in the absence of definite information to the contrary, the conditions of recent years may be considered more similar to those of the present and future than those

---

[1] See W. C. Mitchell, *Business Cycles*, pp. 208–214; and the *Journal of Political Economy*, 21, pp. 515–516; 24, pp. 149, 150 for the data of stock prices.



of remote years. As a consequence of this hypothesis items for recent years would be weighted more heavily than those for remote years, the weights decreasing directly as the distance from the present. Not including a past period is equivalent to giving that period a weight of zero. The weights selected, are of course, arbitrary; that is, they constitute a general judgment concerning relative homogeneity of the items of time series.

The first experiment, then, was to ascertain the effect of weighting. Straight lines were fitted to the wholesale price series for the periods 1879–88, 1879–89, 1879–90, and so on to 1879–1913. In order to give the more recent years greater importance than the earlier ones the years were weighted in arithmetic progression, i. e., 1 to 35, 1879 weighted 1. The segments of the various lines thus determined were plotted in the terminal year.[1] The conclusion, clearly evident and expected, was that this method fails to indicate the secular trend accurately for the period 1896–1913, that is for the period following the decided change in trend of prices. The device of weighting the more recent years more heavily than earlier years gives lines that follow the alternating periods of prosperity and depression too closely to be used as lines of secular trend. It does not appear, moreover, that weighting, or any other mathematical device, can be used as a substitute for direct analysis of homogeneity.

A second experiment was therefore made. Instead of using data for a lengthening period those for a moving ten-year period were taken. Straight lines were fitted to the wholesale price data for the ten-year periods 1879–88, 1880–89, and so on to 1904–13. This method does not determine the trend for the whole period accurately.[2] It is reliable only in case there are no radical changes in direction of trend. Even in cases where the trend does not change decidedly in direction the ten-year period is too short.

The conclusion from these experiments with the wholesale price series of annual items (having a decided downward trend, 1879–96, and a decided upward

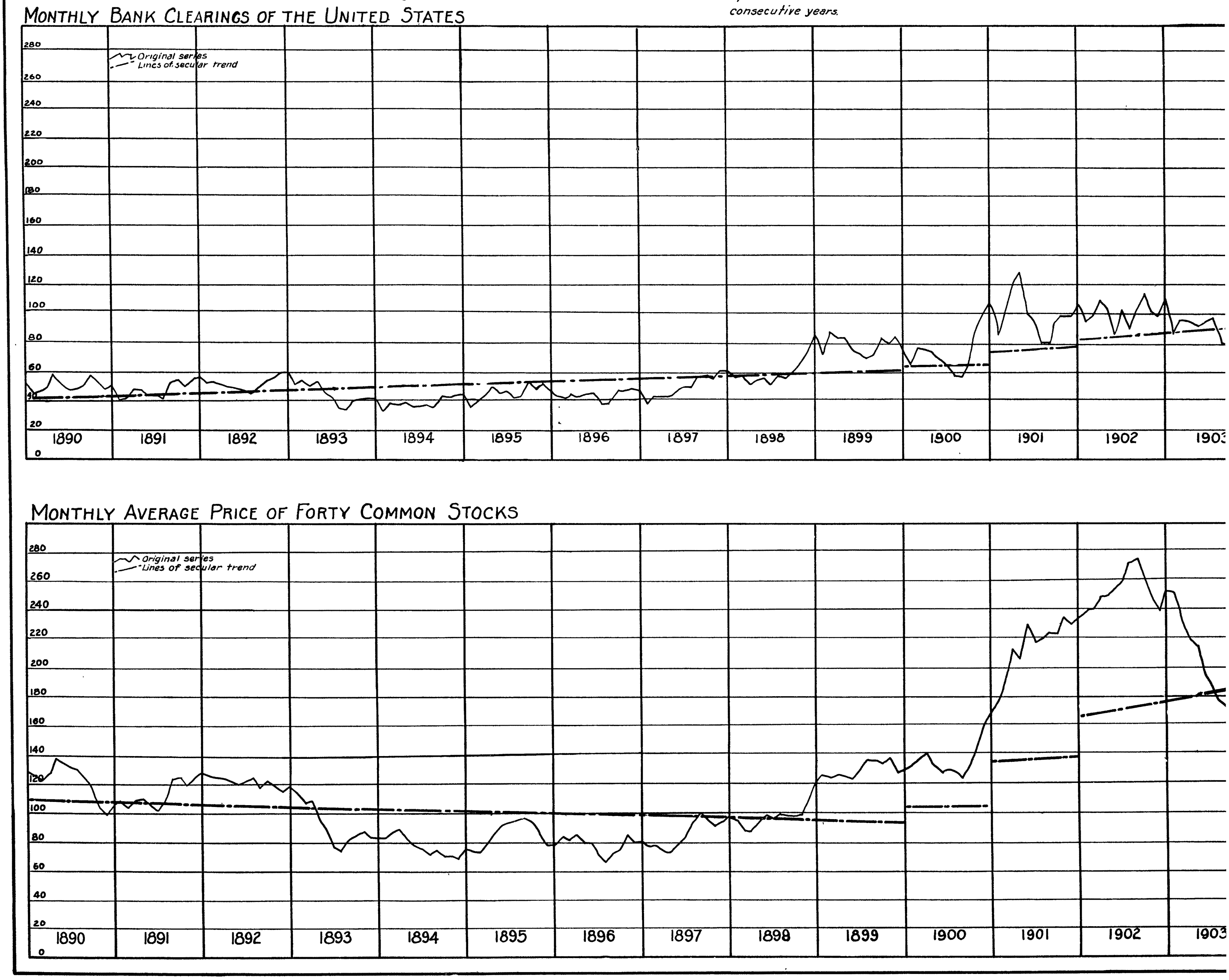
---

[1] The chart has not been reproduced.

[2] The chart on which this judgment is based has not been reproduced.

CHART C

Straight Line Fitted to Monthly Data for Periods 1890-99, 1890-1900, 1890-1901 and so on.
The segments of the consecutive lines thus found are plotted in consecutive years.

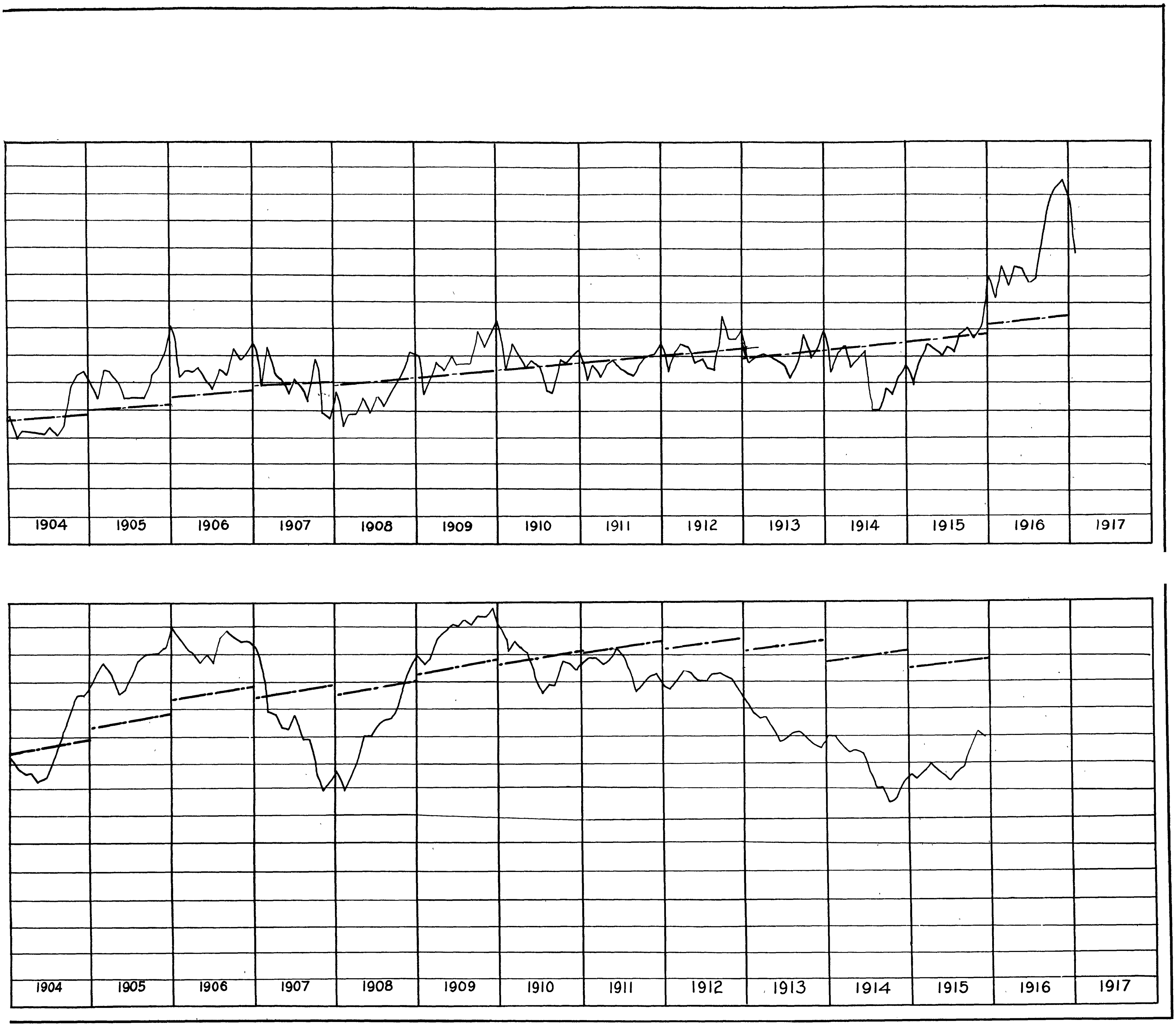


trend, 1897–1913) is unmistakable. It is this: there is a considerable degree of uncertainty in estimating the present secular trend of certain series from data as they come in from year to year. We may, indeed, determine accurately the trend for a fixed past period by use of a curve with a turning point or minimum, such as a parabola or more complex function, but the extension of such a curve is not a reliable method of estimating the trend of the future. To illustrate, Chart B shows that the parabola and the compound interest law both fit the data of per capita annual pig-iron production, 1860–1915, better than a straight line. That fact does not mean, however, that either the parabola or compound interest curve, extended, will picture correctly the future growth of pig-iron production. The same conclusion was reached by inspection of the graphs of parabolas fitted to annual wholesale price indices, and per capita gross receipts of railroads, coal produced, exports, imports, and steel produced for the period 1879–1913.[1]

Does the conclusion mean that any estimate of current and future secular trend is useless? Not at all. It simply means that for certain types of series which have violent major movements covering long periods, in addition to fluctuations corresponding to business prosperity and depression, the prediction of a future trend on the basis of current and past data is subject to a large possible error. The fitting of complex curves to the data will not aid us in obtaining a more reliable estimate than the use of straight lines. Other types of series, however, do not offer the same difficulties. A large number of statistical series, even for an extended

---

[1] The charts have not been reproduced.



period, have a secular trend closely approximating straight lines. For such series use of the straight line method appears feasible. The cases that put the method to test are those in which the trend changes radically in direction.

The conclusions just stated were based on a study of *annual* data. Further experiments were made with *monthly* data, the object being, as in the experiments with annual data, to test the degree of inaccuracy incurred when the straight line of secular trend, computed for a past period, is extended to fit current data. As in the former case, unpromising series were selected, the tests being applied to monthly bank clearings of the United States, 1890–1916, and the average price of forty transportation stocks, 1890–1915 (see Chart C). Assume that we start in 1900 to analyze current data, having complete monthly data for the period 1890–99 in our possession. Straight lines are then fitted to the data in hand. When the data for 1900 come in we correct the line by adding the data for the current year. This gives us a new line of secular trend for the period 1890–1900. Using a similar process we get lines for the periods 1890–1901, 1890–1902, and so on. The segments are plotted in consecutive years. These segments are what we obtain as secular trend by applying the straight line method slavishly during a quarter of a century, the first part of which, 1890–99, is quite unlike the last part, 1900–15. The results are not bad, after all. It is true that for certain years, for instance 1900 to 1903, an estimate of the secular trend of the price of stocks (and to a less degree of clearings) based upon the previous period is unreliable. Estimates based on varying periods, 1890–1900, 1890–1901, 1890–1902, give consecutive lines having ordinates differing by some 10 per cent. Considering the whole period, however, the application of the method would have been distinctly helpful. A more discriminating use of the method might easily have been made to yield better results.

Recognition of the fact that various series have various characteristics with respect to secular trend and that these characteristics should be taken into account in its measurement is one safeguard in measuring secular trend. Another safeguard is the recognition of the facts that trend may be affected either by the accumulation of small influences or by a catastrophe of some kind. In other words, we may not always conclude that the trend of the future will be the same as the trend of the past. The considerations upon which an estimate of trend is based are identical with those at the basis of all plans for the future. The basis of such plans is past experience interpreted in the light of careful observation of current developments. To refuse to estimate or plan for the future because our estimates may be in error is to deny the value of experience and the systematic character of human affairs. With past data available an inexact estimate of secular trend is better than no estimate at all.[1]

We may now set forth our conclusions regarding the different methods that have been examined. The method of curve-fitting is superior to the method of moving averages for measuring secular trend. The determination of a curve or line which pictures the secular trend of a past period, does not determine present or future trend. The presumption that past trend will continue is strong in some cases and weak in others. The estimate of future trend should be influenced by recent tendencies and current items to some degree, yet we should not lightly conclude from short-time fluctuations that secular trend has changed. In a period like 1903–18, for instance, it seems best not to include the war years 1917 and 1918 in fixing the line of secular trend for the present and immediate future. For prices it seems best not to include any of the years 1915 to 1918, during which the price level of the world has risen so enormously. The extension of a past trend is a prophecy. It is impossible to get away from that fact. The important thing is that the exact nature of the prophecy be made unmistakable.

In the application of the method to data for the period 1903–16 all the years were assigned equal weights. A complete explanation is given in Part II in connection with the analyses of individual series.

3. Measurement of Seasonal Fluctuations

The object of the present study of seasonal fluctuations is, first, to determine the existence or non-existence of such fluctuations in various series of monthly data; second, if fluctuations exist, to measure them; and third, to correct the items for seasonal movement. By seasonal movement is meant a consistent variation from one month to the next. Are the items for certain months of the year systematically or regularly different from the items for other months ? If so, there is a seasonal variation.[2]

A consistent variation in a series of monthly totals may be due either (*a*) to a difference in the number of days in the adjacent months, or (*b*) to a difference due to greater activity during exactly equal intervals of time, or to both combined. A variation due to (*a*) may be called an artificial variation, and one due to (*b*) may be called a real seasonal variation. In case we are dealing with items that are not totals, such as prices or interest rates, a consistent variation would not, of course, be due to a difference in the lengths of the months.

A variation in totals due to a difference in the lengths of the months might be corrected by taking daily averages. Certain difficulties, however, would be encountered in getting a correct average. Should we include Sundays or merely take week days ? Should we allow for holidays ? These questions would have to be settled not only for each series but for each year. Moreover, it is not true that the statistical items always refer to calendar months. A " statistical month " ending with the 25th or some other fixed date may have been used. In such a case February would contain 31 days and March 28. In order to make any reliable correction for the varying lengths of the months, the exact usage for each series over a long period of time would have to be ascertained; an error of one day would make a difference of three per cent in the monthly total. In consequence of these uncertainties and the laborious computation required to secure correct daily averages, it was decided to take the data as they are presented and to measure the seasonal variation due to all considerations combined.

That many series of business statistics consisting of monthly or quarterly items present marked seasonal

---

[1] In presenting its Production Index the Federal Reserve Board takes the position that it is useless to attempt to measure secular trend. The *Federal Reserve Bulletin* of December 1918 (p. 1194) contains the following paragraph:

" Objection may be raised to the above method on the ground that no allowance is made for growth — in technical language, that the secular trend is not eliminated. A comparison with conditions existing prior to the outbreak of the war is alone afforded. But reflection will show that this is really no defect at all. For the methods which have been developed for the elimination of the secular trend are all based upon the theory of normal growth, and require that when disturbing factors cease to operate this normal growth shall again be resumed. Who, however, can maintain that the conditions of the past few years may not permanently have changed the direction of the curve which shows this normal trend ? This brings to light the other objection to the attempt to eliminate the secular trend; the attempt is based upon past data only, and hence data which are incomplete and insufficient for the solution of the problem, especially at a time of great change. All that can be done at present is to contrast conditions now prevailing with those prevailing prior to the outbreak of the great war, and thus indicate in their totality the changes which have occurred."

[2] The week appears to be too short an interval for the measurement of seasonal variation. It is not always clear what are corresponding weeks of a series of years. Days are certainly too short intervals to deal with although in some cases it may be profitable to use weekly items. In general it may be said, however, that if a weekly seasonal variation is so evanescent that it does not show in the monthly averages, such variation is not worth while considering for practical purposes. We must be just as careful not to choose too short a period for measuring seasonal variation as to choose too long a period. The month has been adopted here as the most satisfactory unit.